

*Article*

GFP Pattern Recognition in Raman Spectra by Modified VGG Networks for Localisation Tracking in Living Cells

Nungnit Wattanavichean^{1,a,*}, Jirasin Boonchai^{2,4,b}, Sasithon Yodthong^{2,c},
Chakkrit Preuksakarn^{2,d}, Scott C.-H. Huang^{3,e}, and Thattapon Surasak^{4,5,f}

¹ School of Material Sciences and Innovation, Mahidol University, Nakorn Pathom, 73170, Thailand

² Department of Computer Engineering, Kasetsart University, Nakorn Pathom, 73140, Thailand

³ Institute of Communications Engineering, National Tsing Hua University, Hsinchu, 300, Taiwan

⁴ Faculty of Information Technology, Thai-Nichi Institute of Technology, Bangkok, 10250, Thailand

⁵ IOT and Digital Innovation Institute, Digital Economy Promotion Agency, Bangkok, 10900, Thailand

E-mail: ^anungnit.wat@mahidol.ac.th (Corresponding author), ^bjirasin.b@ku.th, ^csasithon.yo@ku.th,

^dchakkrit.p@ku.th, ^echhuang@ee.nthu.edu.tw, ^fs.thattapon@ieee.org

Abstract. The coupling between Raman spectroscopy and green fluorescent protein (GFP) labelling informs chemical compositions at the specific sites. This information leading to study that explain core knowledge of living organism and eventually advance our conventional technique of medical diagnosis. In order to achieve these purposes, the precise interpretation is required. A massive number of Raman/GFP spectra as well as identification of GFP contribution in each spectrum are approaches to achieve those goals. In the paper, CNN is proposed to classify the spectra with and without GFP signal. The dataset of GFP-positive and GFP-negative spectra were created with various size and background color. The feature extraction and classification are conducted with VGG networks. To increase the performance of VGG network, the modified VGG13 and modified VGG19 were designed. These two models extend fully-connected layer from 3 (the original VGG model) to 5 layer for better classification task. Batch normalization is also added at the end of feature extraction units to reduce unpredicted shifting of parameters. The original VGG16, VGG19, and ResNet50 are used as comparison models. The results show that both of our modified VGG models significantly enhances training accuracy of the network comparing to the original VGG. The accuracy of original VGG can be increased when applied pre-trained weight, but the accuracies are yet slightly lower than modified models. Training on ResNet, deeper network, gave the comparable accuracy with our modified models.

Keywords: VGG, machine learning, neural network, ResNet, Raman spectroscopy, fluorescence, living cells.

ENGINEERING JOURNAL Volume 25 Issue 2

Received 30 July 2020

Accepted 29 January 2021

Published 28 February 2021

Online at <https://engj.org/>

DOI:10.4186/ej.2021.25.2.151

1. Introduction

Raman spectroscopy is the non-destructive technique based on principal interaction between light and electron of matter [1], [2]. This technique shows the inelastic scattering of vibrational modes which specific to particular molecules. The different in vibration refers to molecular properties, stages, or chemical compositions, and they show obviously in the spectrum. According to its simplicity and non-destruction, Raman spectroscopy is widely used in many fields of science such as Chemistry, Physics, Biology, and medical [3], [4], [5], [6], [7], [8], [9], [10] especially for the living systems. In order to study living system, traceability is the critical concern. Green fluorescent protein (GFP) is the typical protein which can emit green fluorescent and visible to bare eyes [11]. As this straightforwardness, it is extensively used for genetically targeting organelles or other proteins in living cell to visualize only interested components. The most advantage of using GFP is because its emission not only observe qualitatively via human eyes, but can also measure quantitatively with spectrometer as GFP spectrum. Combination between Raman spectroscopy and GFP tagging elevates the study of dynamic system in living organisms since the information about chemical composition and location are presented on a single spectrum [12]. The component analysis of living system is very powerful information for further fundamental understanding about life. When comes into medical field, it can developed to be more sensitive diagnosis technique in the future. However, to classify only targeted spectrum is a handful process in which the spectrum is selected one by one with human eyes. A mathematical program is used, but the parameters are need to adjust by human. There two conventional approaches may leads to misinterpretation of Raman/GFP spectra. Therefore, in this paper we proposed an idea to classify Raman/GFP spectra by incorporated with artificial intelligence.

Artificial Intelligence (AI) is a branch of computer science which aims to simulate computers to learn, decide and solve a problem like a human [13]. The subfield of AI such as Machine Learning is a core of AI that makes computers more intelligent by feeding data, thus machines can perform automatically learning itself without supervision required. AI becomes a part of Industry technology, especially in the manufactory [14]. Many applications are also applied with machine learning such as image processing, natural language processing, and robotics [15]. Nowadays, AI is widely used in image classification with deep learning algorithms which is one of the best algorithms in machine learning. Its neural network is inspired by the working of human brain [13]. One of neural networks model that good in image classification is convolutional neural network.

Convolutional Neural Networks (CNNs or ConvNets) are the networks which have considerable success in image recognition with large-sized dataset [16], [17], [18], [19]. Due to the achievement in classification, CNNs became one of the most popular neural networks and played a

huge role in many works over the last few years, especially in the medical field. CNNs have been applied in circumstance of MRS which analyzed biochemical changes in the brain with waveforms [20]. It was also used for identification of pathogens in relevant biofluids which is more accurate. It can improve patient outcomes and reduce healthcare costs [21]. The extract image feature of CNN can help in spectral prediction [14]. It is used as an approach to detect the artefact in wideband Near-Infrared Spectroscopy (NIRS) [22]. In the field of Raman spectroscopy, its classification achieved significantly high accuracy [23].

Deep convolutional networks showed the competence in image classification which can extract more specific and complex features as well as better classification performance [14], [19], [24], [25]. The architecture that we are interested in are VGG16 and VGG19. VGG are the models which are easy to understand and uncomplicated to adjust the layer. Moreover, these models can apply and fine-tune to fit small datasets [26], [27], [28], [29].

In this paper, we modified our models based on VGG architectures to suit the GFP dataset which is not very big. Our two modified models have a difference in the depth of convolutional layers, and we also add batch normalization to fix the problem from the change in the distribution of layers' inputs cause the parameters of the previous layers change also known this problem as 'Internal Covariate Shift' which will slows down the training and make the model hard to train [26], [30]. By shrinking the dimension of fully-connected layers, the parameters are significantly reduced. More fully-connected layers are added to improve accuracy in classification tasks. The transfer learning which is the use of pre-trained weight for classification steps is applied [28], [29], [31], [32]. It is used to compare the accuracy with our modified VGG networks. Residual neural network (ResNet) is also used as a reference network since it has a great performance, and it commonly used with Raman spectra data [33] and solved the model degradation problem on deep networks [34].

This paper consists of 6 sections. Section 2 describes the system overview, the details of data pre-processing and our dataset that we use in the training step. The neural network architecture is presented in section 3. Section 4 show the detail of network training and the results performance of our models in GFP classification. Finally, the discussion and conclusion are described in Section 5 and section 6, respectively.

2. Proposed Method

2.1. System Overview

The two main processes in our system are image generation and model training, see Fig. 1. At the beginning, raw data from excel file were converted spectral images with the User Interface (UI). At the UI, we can import an external file and select image size and background color

for generating spectral image. The preprocessing step is added to cut out obvious defect from images. Then, the training set of spectral image is introduced to our network for training. The models we used for training are our modified VGG13, modified VGG19, original VGG16, original VGG19, and ResNet to compare the accuracy. All CNN networks are implemented with Keras and Tensorflow. The testing set is used after training step to verify network accuracy. The output is shown as probability value between 0-1 which indicates the prediction accuracy to be GFP-positive and GFP-negative (see section 2-2.3). A program has been developed to store classified data and saved in separated folders.

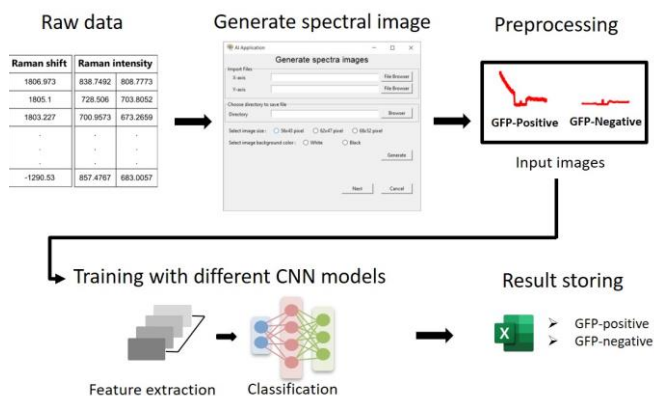


Fig. 1. The overview of our system: Raman/GFP raw data collection, spectral image generation, data pre-process, training in CNN network, and classification result storage.

2.2. Spectral Image Generation

The raw data were obtained as Raman/GFP spectra of *Shizosaccharomyces pombe* mitochondria from Prof. Hiro-o Hamaguchi, Department of Applied Chemistry, National Chiao Tung University, Taiwan [12]. The image generation process from raw data are demonstrated in the early part of Fig. 1. In brief, Raman shift and Raman intensity are assigned to arrays in NumPy library for x- and y-axis respectively. The program for image generation is written in Python. The Panda library is used to help reading raw data from excel file. All the images are preprocessed by removing cosmic rays from the spectrum. The spectral images are generated as 3 sizes, 56x43, 62x47, and 68x52. All the spectra images are converted to RGB pixels at the value of 255,255,255 for white background and 0,0,0 for black background. All images are eventually saved to .png file.

The user interface is created to generate spectral images by importing external raw data files. The images are then automatically labelled either GFP-positive or GFP-negative for training process in the next step. The UI can generate the dataset with different size and background color as shown in Fig. 2. In the final step, the prediction result from training in neural network is obtained. This program can compare the result with raw

data, and separately store GFP-positive and GFP-negative data in different folders for further analysis.

2.3. Dataset

All datasets were created by the same raw data of 30,000 spectra with other specifications below:

- Size 56x43 pixel, white background color
- Size 62x47 pixel, white background color
- Size 68x52 pixel, white background color
- Size 56x43 pixel, black background color
- Size 62x47 pixel, black background color
- Size 68x52 pixel, black background color

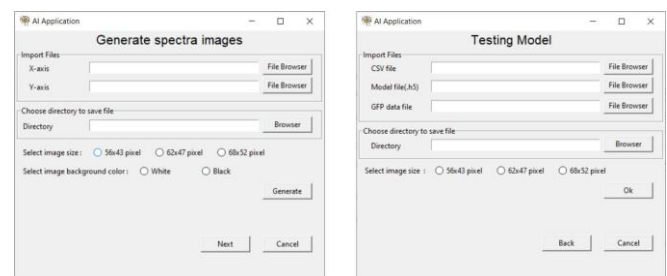


Fig. 2. The user interface to generate spectral images to preferred specification.

All datasets were separated to training and testing sets in the ratio of 26,000 to 4000 spectral images. The ratio between GFP-positive and GFP-negative is 50:50 for both training and testing sets. The example of spectral images are shown in Fig. 3. The GFP-positive label is the spectra which recognized as GFP-containing spectra and the GFP-negative label referred to non-GFP containing spectra as examples in Fig. 3. The difference of GFP-positive and GFP-negative images can be seen from low x-value region (half-left of the generated image). GFP-positive exhibits high y-value at low x-value region while GFP-negative image shows y-value almost zero at low x-value.

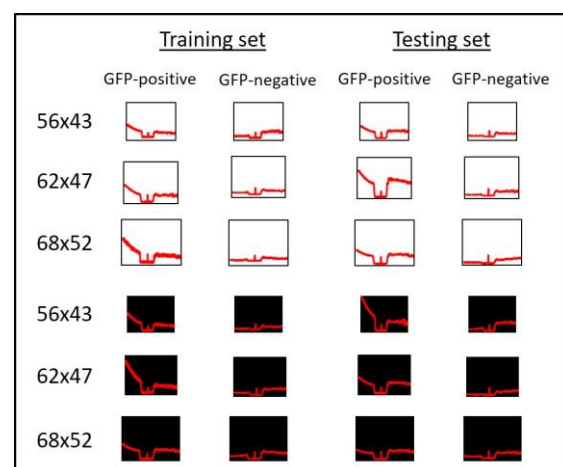


Fig. 3. The illustration of GFP-positive and GFP-negative images in each dataset. The image size and background color are indicated on the left and top of the images, respectively.

3. Neural Network Architecture

3.1. Original VGG16 and VGG19

According to performance of deep convolutional neural networks in large-scale image recognition [19], two outstanding models are VGG16 and VGG19 which are the first two models in Fig. 4. The main components of these two models are convolutional layers, max-pooling layers, fully-connected layers and softmax. VGG16 consists of 13 layers of convolutional but VGG19 comprises 16 convolutional layers. Both of them have 5 max-pooling layers, 3 fully-connected layers and softmax in the final layer. All convolutional layers of these two models are divided into 5 feature extraction units. The number of filters in the first feature extraction unit starts from 64 and then increases by a factor of 2 in each group until 512 in the final unit. The stride and padding are fixed to 1 in all convolutional layers. Max-pooling after convolutional layers in each group is performed over a 2x2 pixel window with stride 2. Three Fully-connected layers, the first two performs 4096 channels in each layer and the third has 2 for 2 classes of our GFP classification. Softmax is added at the end of classification step to adjust classification result and present as a probability.

3.2. Our Modified VGG16 and VGG19

Our VGG models are developed to enhance classification tasks of neural networks. The number of fully-connected layers are increased from 3 layers (original VGG network) to 5 layers. In addition, batch normalization is also added at the end of every feature extraction unit (convolution layer and max-pooling layer) to help shifting of internal covariate and improving the training process [26], [30]. Modified VGG13 model is adapted from the original VGG16. It consists of 3 sets of feature extraction units, 5 fully-connected layers, and softmax. The increase of fully-connected layers means the longer training period. Therefore, this modified model is reduced feature extraction units from 5 iterations to 3 iterations to compensate time consumption. Another modified model is modified VGG19. This model consists of 5 sets of feature extraction units with batch normalization, 5 fully-connected layers, and softmax. The model is generally similar to the original VGG19 except the number of convolutional layers in each feature extraction unit are different (Fig. 4). The fully-connected layers of both of our modified VGG models are adjusted to classify GFP-positive and GFP-negative tasks. The first fully-connected layer consists of 2048 channels. The channels are decreased by a factor of 2 in the next layer until reaching the fourth layer, which is 256 channels. The final fully-connected layer performs 2-way GFP classification. The overall architecture of two modified VGG models are illustrated in Fig. 4.

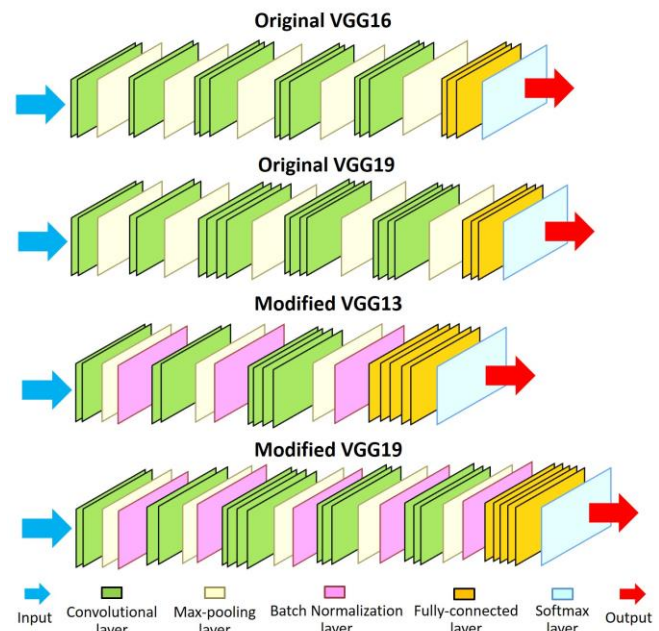


Fig. 4. Overall network architectures, compare between original VGG and modified VGG networks in the depth of CNN. The layers compose of convolutional layer, batch normalisation, max-pooling, and fully-connected layers. From the top: The original VGG16, the original VGG19, the modified VGG13, and the modified VGG19.

3.3. ResNet50

In the field of neural networks, the developers believe in the fact that the more convolutional layers refer to the better neural network performance. However, a number of literatures said too many convolutional layers can also cause efficiency drop of the model.

ResNet or Residual Network is one example of a very deep network. It can raise the accuracy by increasing the depth of network [34]. ResNet50 model uses the modified building block called a bottleneck design which is a stack of three layers consisting of 1x1, 3x3, and 1x1 convolutions. It also has batch normalization after each convolutional layer and before the activation. ResNet50 architecture contains 50 layers including a single fully-connected layer at the terminal followed by softmax. In this paper, a fully-connected layer is adjusted to suit our data. Therefore, the channels are changed from 1000 to 2 channels according to 2 classes of GFP. The special part of this network is shortcut connection which can skip some unimportant convolutional layers. It allows only the essential data passes through the next layer. In our work, ResNet with 50 layers depth is used to compare the accuracy and latency with our modified and original VGG models which contain less layers.

4. Experiments and Results

4.1. Network Training

Original VGG16, VGG19 and ResNet50 models are the pre-trained models from Keras which is fast and easy for fine-tuning. The modification is done to the last layer of fully-connected layer to fit with our GFP classification. These previous three models are used as reference to compare the performance with our modified VGG networks. In order to get higher accuracy, the ImageNet weight is borrowed from pre-trained model in Keras. ImageNet weight is obtained from training of a very large-scale dataset. We believe that the pre-trained weight can improve the training accuracy. The condition of using pre-trained weight (ImageNet) and without pre-trained weight are compared to find the optimized model for our datasets. The training is conducted with 4 VGG models that we mentioned earlier (2 original VGG and 2 modified VGG). Another important factor for the training is the dataset. The size and background color are varied as described in section 2-2.3. The size has been generated to 3 sizes consisting of 56x43, 62x47 and 68x52 pixels. In each size of GFP spectra images consist of background colors, which is white and black. In addition, the two original VGG networks are also trained without weights and weights from ImageNet for comparison. The comparison of VGG models with ResNet50 is performed eventually to see the effect of network depth to the classification achievement.

background datasets are used for further comparison of the accuracy between modified VGG and the original VGG networks.

The accuracy of the modified VGG13 model is around 4-10% higher than the original VGG16 as well as giving lower SD values. We want to note that the original VGG16 networks which trained with pre-trained weight can perform very good classification regardless white or black background. The training accuracy is around 82 - 84% for both colors. However, when comparing the accuracy with our modified VGG13 model, the accuracy of the original VGG16 with pre-trained weight is slightly lower. The accuracy trend of the modified VGG19 and the original VGG19 also have consistent responses with that of modified VGG13 and the original VGG16. The highest accuracy from the modified model and the original VGG19 achieved from 68x52 pixel image with black background color dataset. The training accuracy of the modified VGG19 and the original VGG19 are $85.84 \pm 0.41\%$ and $76.04 \pm 1.03\%$, respectively. Due to these accuracies, we can imply that our modified VGG19 model works better than the original VGG19 model. Even the original VGG19 with pre-trained weight which gave a high accuracy, yet still lower than the modified VGG19. According to the result from Table 1, the size of the input image can be considered as negligible factor

Table 1. Training accuracy from our modified VGG19 network compared to original VGG19 and original VGG19 with weight from ImageNet.

	Image size / pixel	Modified VGG		Original VGG		Original VGG (ImageNet)	
		White	Black	White	Black	White	Black
VGG16*	56x43	65.82 ± 16.02	85.35 ± 0.78	60.05 ± 4.67	81.77 ± 3.64	83.83 ± 0.36	84.18 ± 0.50
	62x47	66.19 ± 12.67	85.93 ± 1.16	55.58 ± 3.55	79.96 ± 2.74	83.51 ± 0.25	82.50 ± 1.23
	68x52	71.65 ± 18.84	85.44 ± 0.78	59.36 ± 5.56	75.53 ± 1.23	84.03 ± 0.44	83.41 ± 0.67
VGG19	56x43	69.40 ± 16.84	83.80 ± 0.48	54.64 ± 2.04	55.38 ± 7.32	85.57 ± 0.32	84.47 ± 0.10
	62x47	53.02 ± 5.23	85.64 ± 0.63	51.61 ± 2.79	62.24 ± 8.20	80.73 ± 1.37	83.74 ± 0.30
	68x52	50.11 ± 0.13	85.84 ± 0.41	52.16 ± 3.54	76.04 ± 1.03	84.55 ± 0.42	84.70 ± 0.13

*Refers to modified VGG13 or the original VGG16

4.2. Results

The training accuracy of modified VGG and original VGG models are shown in Table 1. The training accuracy of our modified VGG13 is compared with the original VGG16 model, top three rows. The modified VGG19 is compared with the original VGG19, bottom three rows. For two original VGG models, the networks were trained with pre-trained weight (ImageNet) and without pre-trained weight, middle and right column of Table 1 respectively. The different datasets which are size and background color also indicated on the row and column heads. The training results of the modified VGG and original VGG without pre-trained weights show the accuracy of white background color in every image size is lower than the case that trained with black background color. The standard deviation (SD) are also higher in white background datasets which refers to the more stable network with black background training. Therefore, black

since the accuracies are slightly changed. In contrast, background color plays a significant role in improving accuracy. Black background color not only performs better classification, but it also reduces the time and memory consumption of training process.

In order to demonstrate the performance of our modified VGG network, a confusion matrix was created, see Fig. 5. The dataset of 68x52 pixel with black background is tested with the modified VGG19. This condition gave the highest accuracy for all VGG19 models, average accuracy $85.84 \pm 0.41\%$. The confusion matrix shows the classification output of our modified VGG19 model. The output is presented by number of images which classified as GFP-positive and GFP-negative in four quadrants. The agreement between categories in x- and y-axis is counted as correct classification. This test set contains 4000 spectral images. They are labeled as GFP-positive and GFP-negative by 50:50. The confusion matrix reveals that our model can predict the correct GFP-

positive and GFP-negative at 1799 and 1653 out of 2000, respectively. The total fault prediction is 548 images, which converted to 86.30% accuracy.

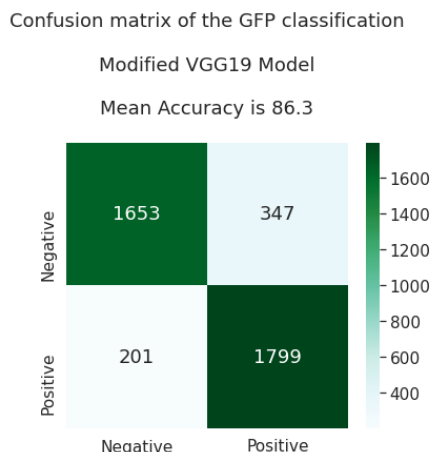


Fig. 5. Confusion matrix of 68x52 pixel with black background color dataset trained with the modified VGG19 network.

After we compared between the modified and the original VGG networks, we also did the comparison with the ResNet50 model. Figure 6 shows the accuracy of every network that we compared by using one of the best conditions dataset which is 68x52 pixel with black background. The accuracies of ResNet50 with and without ImageNet weights pre-trained have similar results around 85%, but the ResNet50 which initialized with ImageNet weights has slightly higher accuracy as well as lower SD. Besides, the ResNet50 gave a greater performance than the original VGG16 and VGG19. The initial weights from ImageNet can improve training accuracy and lower SD as found in the original VGG and ResNet50. However, without the initial weight, our modified VGG model gave the best performance in accuracy and SD values.

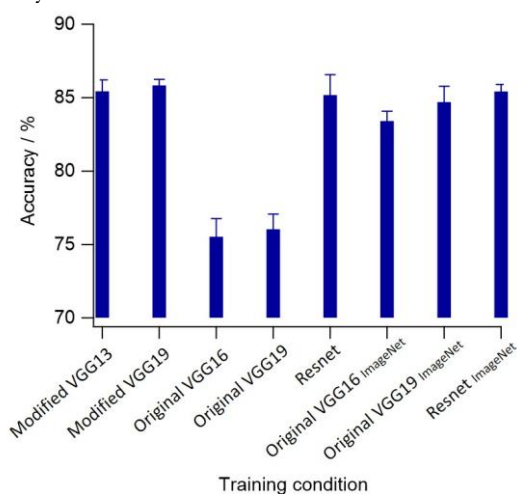


Fig. 6. Training accuracy comparison of the modified VGG13, modified VGG19, original VGG16, original VGG19, and ResNet50. The pre-trained weight from ImageNet are applied for original VGG16, original VGG19, and ResNet50.

The convergence rate of individual models also evaluated as illustrated in Fig. 7. The original VGG19 shows the slowest convergence, followed by the original VGG16. These 2 models converge better when applied with pre-trained weights. ResNet50 and our modified VGG models have faster converge rate than the original VGG. The reason for the fast convergence is from batch normalization [26], [30] in our model and ResNet50. However, our modified networks have more fluctuations than other networks. Apart from the convergence rate, the training time consumption also observed. The modified networks used the time around 20-50s per epoch which are close to the originals. ResNet50, on the other hand, spent around 60-200s per epoch due to deeper layer of feature extraction. The shorter training period of our modified models also caused from massive reductions of parameters. The parameters in the original VGG network are 3 times higher and twice for the ResNet50.

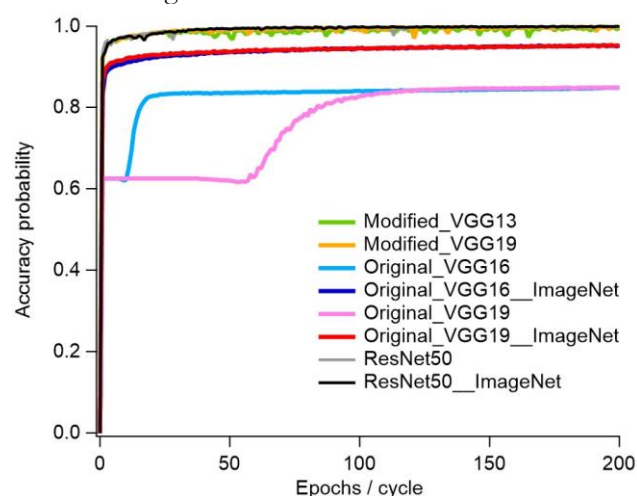


Fig. 7. Convergence rate of the modified VGG13, modified VGG19, original VGG16, original VGG19, and ResNet50. The pre-trained weight from ImageNet are applied for original VGG16, original VGG19, and ResNet50.

5. Discussion

The image sizes in the dataset are separated into 3 sizes and 2 background colors which are 68x52, 62x47 and 56x43 pixels with white and black. Our modified models and the original VGG models without using ImageNet pre-trained weights in Table 1 shows that the black background dataset outperformed in classification than the white color. On the other hand, different sizes did not give huge impact to the accuracy in our modified models but affected the original networks. The original VGG models with pre-trained weights gave the low accuracy, even the dataset was changed from white to black color. Therefore, the use of VGG network without any modification of convolutional and fully-connected layers is not appropriate to Raman/GFP dataset. ResNet50 was used for comparison as shown in Fig. 6 by using the finest condition that always gave a high accuracy and low SD. The bar chart shows the performances of the ResNet50

model with and without applying pre-trained weight are very similar. Nonetheless, the original VGG16 and VGG19 performed exceedingly better when ImageNet weights were used. Even though the original VGG16, original VGG19 and ResNet50 pre-trained model gave the high level accuracy, our modified VGG models still had the highest accomplishment. Deep convolutional neural networks are used in various kinds of spectrum patterns to improve the work efficiency. In Near-Infrared Spectroscopy Sensors, CNN with parameter initialization can improve data analysis. MSRA, Gaussian distribution of specific variance, was used as initial weight [35]. They received prediction accuracy around 93% from the training of 990,000 cycles. The highest accuracy of 93.13% was obtained when setting the best conditions of learning rate, CNN layers, activation function and network optimization evaluated with 2674 samples in the testing set. For Raman spectra classification, multilayer perceptron (MLP) algorithm has been implemented to classify mineral spectra [36], the classification performance of their neural networks showed 83% and 80.4% accuracy, when trained with Raman spectral data from their mineral collection and the RRUFF mineral library, respectively. Our work in Raman/GFP spectra classification reveal the highest mean accuracy around 85-86% from the modified VGG networks by training with Raman/GFP dataset less than 200 rounds. Compared to other works that we mentioned previously, our deep convolutional networks work better with Raman spectra using much shorter training periods as well as obtain higher accuracy. The important components in our modified model which gain accuracy and help time-consuming processes are batch normalization and fully-connected layers. Batch normalization aids each layer of the networks have a normal distribution avoid Internal Covariate Shift problem and also made our networks converge faster [26], [30]. Fully-connected layers can impact the classification performance [22], the number of neurons in FC layers as well as more number of FC layers can obtain better performance [37]. Finally, the data pre-processing is also a crucial step because the color and size or quality of images that are used can affect the training and learning capability of the networks [38], [39].

6. Conclusion

The VGG network was used to assist classification GFP appearance on the Raman/GFP spectra for the first time. The modified VGG13 and VGG19 were adapted from the original VGG16 and VGG19 by reducing convolution layer, adding Batch normalize and increase fully-connected layers to improve the training performance. Datasets are Raman/GFP spectral images which various sizes, 56x43, 62x47, and 68x52, and background colors of black and white. The original VGG networks were initialized with ImageNet pre-trained weight to evaluate the change to accuracy. It can improve the accuracy in the order of 10-20% for the original VGG networks. The training accuracy of our modified models

in both models provided higher accuracy than original VGG19 and original VGG16, even the pre-trained weight was applied. The best VGG model in classification is our modified VGG19, achieving 85.84% accuracy from 68x52 pixel image with black background dataset. Resnet50 model also used to compare with our VGG model. It gave the similar accuracy with modified VGG19 which is around 85% for both with and without pre-trained weight. However, ResNet50 consumed longer time and higher memory for training process. Therefore, with the factors of training accuracy, time and resource management, our modified VGG models are proof to be the best network to train and classify GFP-positive and GFP-negative spectral images. The classified spectral images were eventually stored in separated folder for further Raman spectral analysis.

Acknowledgement

The authors would like to thank Prof. Hiro-o Hamaguchi, National Chiao Tung University, Taiwan for Raman spectroscopic experiment. *S. pombe* yeast strain was obtained from laboratory of Prof. Makoto Kawamukai and Dr. Ikuhisa Nishida with collaboration of Prof. Tatsuyuki Yamamoto, Shimane University, Japan.

References

- [1] O. I. Olubiyi, F.-K. Lu, D. Calligaris, F. A. Jolesz, and N. Y. Agar, "Advances in molecular imaging for surgery," in *Image-Guided Neurosurgery*. Boston: Academic Press, 2015, ch. 17, pp. 407–439.
- [2] P. Rostron and D. Gerber, "Raman spectroscopy, a review," *International Journal of Engineering and Technical Research*, vol. 6, pp. 50–64, 2016.
- [3] N. Wattanavichian, E. Casey, R. J. Nichols, and H. Arnolds, "Discrimination between hydrogen bonding and protonation in the spectra of a surface-enhanced Raman sensor," *Physical Chemistry Chemical Physics*, vol. 20, no. 2, pp. 866–871, 2018.
- [4] N. Wattanavichian, M. Gilby, R. J. Nichols, and H. Arnolds, "Detection of metal–molecule–metal junction formation by surface enhanced Raman spectroscopy," *Analytical Chemistry*, vol. 91, no. 4, pp. 2644–2651, 2019.
- [5] X. Wang, J. Shen, and Q. Pan, "Raman spectroscopy of sol–gel derived titanium oxide thin films," *Journal of Raman Spectroscopy*, vol. 42, no. 7, pp. 1578–1582, 2011.
- [6] F. T. Prochaska and L. Andrews, "Vibration–rotational and pure rotational laser–Raman spectra of h₂, d₂, and hd in matrices at 12 k," *The Journal of Chemical Physics*, vol. 67, no. 3, pp. 1139–1143, 1977.
- [7] H. J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, K. Esmonde-White, N. J. Fullwood, B. Gardner, P. L. Martin-Hirsch, M. J. Walsh, M. R. Mcainsh, N. Stone, and F. L. Martin, "Using Raman spectroscopy to characterize biological materials," *Nature Protocols*, vol. 11, no. 4, pp. 664–687, 2016.

- [8] N. Kuhar, S. Sil, T. Verma, and S. Umaphathy, "Challenges in application of Raman spectroscopy to biology and materials," *RSC Advances*, vol. 8, no. 46, pp. 25 888–25 908, 2018.
- [9] K. Kong, C. Kendall, N. Stone, and I. Nottingher, "Raman spectroscopy for medical diagnostics — from in-vitro biofluid assays to in-vivo cancer detection," *Advanced Drug Delivery Reviews*, vol. 89, pp. 121–134, 2015.
- [10] I. Pence and A. Mahadevan-Jansen, "Clinical instrumentation and applications of Raman spectroscopy," *Chemical Society Reviews*, vol. 45, no. 7, pp. 1958–1979, 2016.
- [11] S. J. Remington, "Green fluorescent protein: A perspective," *Protein Science: A Publication of the Protein Society*, vol. 20, no. 9, pp. 1509–1519, 2011.
- [12] N. Wattanavichian, I. Nishida, M. Ando, M. Kawamukai, T. Yamamoto, and H.-O. Hamaguchi, "Organelle specific simultaneous Raman/green fluorescence protein microspectroscopy for living cell physicochemical studies," *Journal of Biophotonics*, vol. 13, no. 4, p. e201960163, 2020.
- [13] G. Ciaburro and B. Venkateswaran, *Neural Networks with R: Smart models using CNN, RNN, Deep Learning, and Artificial Intelligence Principles*. Packt Publishing, 2017.
- [14] J. Lee, Q. Duan, S. Bi, R. Luo, Y. Lian, H. Liu, R. Tian, J. Chen, G. Ma, J. Gao, and Z. Xu, "Machine learning promoting extreme simplification of spectroscopy equipment," *arXiv preprint*, arXiv:1808.03679, 2018.
- [15] V. Jeyamani, J. Ashok and S. Suppiah, "A review on significance of sub fields in artificial intelligence," *International Journal of Latest Trends in Engineering and Technology*, vol. 6, no. 3, pp. 542-548, 2016.
- [16] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097-1105, 2012.
- [17] M. Zeiler and R. Fergus, "Visualizing and understanding convolutional neural networks," in *European Conference on Computer Vision*, 2013, vol. 8689, pp. 818-833.
- [18] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. Lecun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," in *International Conference on Learning Representations (ICLR) (Banff)*, 2013.
- [19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv 1409.1556*, 2014.
- [20] S. Gurbani, E. Schreibmann, A. Maudsley, J. Cordova, B. Soher, H. Poptani, G. Verma, P. Barker, H. Shim, and L. Cooper, "A convolutional neural network to filter artifacts in spectroscopic MRI," *Magnetic Resonance in Medicine*, vol. 80, no. 5, pp. 1765-1775, 2018.
- [21] C.-S. Ho, N. Jean, C. Hogan, L. Blackmon, S. Jeffrey, M. Holodniy, N. Banaei, A. Saleh, S. Ermon, and J. Dionne, "Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning," *Nature Communications*, vol. 10, no. 1, pp. 1-8, 2019.
- [22] D. Jacobson, "A machine learning approach to artefact detection in broadband near-infrared spectroscopy (NIRS)," summer project, Department of Medical Physics & Biomedical Engineering, University College London, 2018.
- [23] J. Liu, M. Osadchy, L. Ashton, M. Foster, C. Solomon, and S. Gibson, "Deep convolutional neural networks for Raman spectrum recognition: A unified solution," *The Analyst*, vol. 142, no. 21, pp. 4067-4074, 2017.
- [24] D. Ciresan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Flexible, high performance convolutional neural networks for image classification," in *International Joint Conference on Artificial Intelligence*, 2011, pp. 1237-1242.
- [25] W. Yu, K. Yang, Y. Bai, H. Yao, and Y. Rui, "Visualizing and comparing convolutional neural networks," *arXiv preprint*, arXiv:1412.6631, 2014.
- [26] S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," in *2015 3rd LAPR Asian Conference on Pattern Recognition (ACPR)*, 2015, pp. 730-734.
- [27] M. Mateen, J. Wen, D. Nasrullah, S. Song, and Z. Huang, "Fundus image classification using VGG-19 architecture with PCA and SVD," *Symmetry*, vol. 11, no. 1, 2018.
- [28] K. B. Ahmed and A. Jelodar, "Fine-tuning VGG neural network for fine-grained state recognition of food images," *arXiv preprint*, arXiv:1809.09529, 2018.
- [29] M. Shu, "Deep learning for image classification on very small datasets using transfer learning," master's thesis, Iowa State University, 2019.
- [30] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France, 2015, vol. 37.
- [31] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43-76, 2021.
- [32] S. Tammina, "Transfer learning using VGG-16 with deep convolutional neural network for classifying images," *International Journal of Scientific and Research Publications (IJSRP)*, vol. 9, no. 10, pp. 143-150, 2019.
- [33] X. Chen, L. Xie, Y. He, T. Guan, X. Zhou, B. Wang, F. Guangxia, H. Yu, and Y. Ji, "Fast and accurate decoding the Raman spectra encoded suspension array with deep learning," *The Analyst*, vol. 144, no. 14, pp. 4312-4319, 2019.
- [34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770-778.

- [35] Z. Wang, Tian, Y. Jx, Z. Zhu, J. Jiang, and Y. Cai, "Improved deep CNN with parameter initialization for data analysis of near-infrared spectroscopy sensors," *Sensors*, vol. 20, p. 874, 2020.
- [36] S. Ishikawa and V. Gulick, "An automated classification of mineral spectra," in *Lunar and Planetary Science Conference*, 2013, vol. 1719, p. 3085.
- [37] S. S. Basha, S. R. Dubey, V. Pulabaigari, and S. Mukherjee, "Impact of fully connected layers on performance of convolutional neural networks for image classification," *Neurocomputing*, vol. 378, pp. 112-119, 2020.
- [38] S. Bianco, C. Cusano, P. Napoletano, and R. Schettini, "Improving CNN-based texture classification by color balancing," *Journal of Imaging*, vol. 3, no. 33, 2017.
- [39] S. Dodge and L. Karam, "Understanding how image quality affects deep neural networks," in *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, 2016, pp. 1-6.



Nungnit Wattanavichean was born in Bangkok, Thailand on 24th January 1990. She received her B.Sc. degree in Biotechnology at Faculty of Science, Mahidol University, Thailand. She had completed her Ph.D. program in Department of Chemistry, University of Liverpool, UK in 2016. Her thesis was about Raman spectroscopy of molecular electronic junctions.

In 2017-2019, she became a postdoctoral researcher in Department of Applied Chemistry, National Chiao Tung University, Taiwan. She is currently working as a lecturer in School of Materials Science and Innovation, Mahidol University, Thailand.

Dr. Wattanavichean research areas are about Raman spectroscopy of electronic junctions and living organisms, biomaterials, and cosmetic science.



Jirasin Boonchai was born in Kanchanaburi, Thailand. He received his B.Eng. degree (first-class honors) in Computer and Electronics Engineering from Kasetsart University (Nakhon Pathom, Thailand) in mid-2020, and currently joined Faculty of Information Technology at Thai-Nichi Institute of Technology (Bangkok, Thailand) as a master's degree student (with scholarship) since mid-2020. He is now working with the Digital Service and Platform Promotion Department of the Digital Economy Promotion Agency (Bangkok, Thailand) as an officer.



Sasithon Yodthong received the B.Eng. degree from the Department Computer and Electronics Engineering of Kasetsart University Kamphaengsaen Campus (Nakhon Pathom, Thailand) in mid-2020 and she is currently joined been Bang-up Product Co., Ltd. (Bangkok, Thailand) as a Programmer since late-2020.



Chakkrit Preuksakarn is an Assistant Professor in the Department of Computer Engineering at Kasetsart University, Kamphaengsaen Campus, Thailand. He completed his Ph.D. at Montpellier II University, France. He has collaborated actively with researchers in several other disciplines of engineering and science. His research interests include programming language, logical frameworks, web services, and artificial intelligence.



Scott C.-H. Huang received his B.S. degree in Mathematics from National Taiwan University (Taiwan) in 1998, and his Ph.D. degree from the University of Minnesota Twin Cities (USA) in 2004. He is currently an Associate Professor in the Department of Electrical Engineering and Institute of Communication Engineering, National Tsing Hua University (NTHU). Prior to joining NTHU, Dr. Huang was at City University of Hong Kong from 2005 to 2010.

Dr. Huang's research interests include FinTech, Blockchain-based Applications & Services, Information Security, Network Security, Wireless Networking, Internet of Things, Digital Signal Processing and Error Correcting Codes.



Thattapon Surasak received his B.Eng. degree in Computer Engineering from Kasetsart University, Nakhon Pathom, Thailand in 2014, M.Sc. degree with distinction in Telecommunications Engineering from the University of Sunderland, Sunderland, United Kingdom in 2016, and his Ph.D. in Communications Engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2020.

Dr. Thattapon Surasak joined the Digital Economy Promotion Agency, Bangkok, Thailand, in October 2020, where he is now a Senior Team Leader of the IOT and Digital Innovation Institute. He is a Reviewer for IET Smart Cities, ECTI Transactions on Electrical Engineering, Electronics, and Communications, and he is also a Mentor for the Innovative Incubation Center of National Tsing Hua University. His research interests include Blockchain-based Applications & Services, Distributed Database Technologies for Blockchain, Data Analytics, Wireless Networks and Network Security.